Methods for simulation of biological and/or chemical reaction pathway, biomolecules and nano-molecular systems

Field of the invention

The present invention relates to Method for simulation of biological pathways, chemical reaction pathways, biomolecules, and nano-molecular-machines by using electronic circuits.

This general field is known as "Molecular Biology" (MB), "Cell Biology" (CB), "Nano-technology" (NT), "Chemistry" (CH). When used for analysis of disease processes, this field is referred to as "Pathology". When used for the examination of the effect of pharmaceutical agents on biological systems, this field is referred to as "Drug evaluation" (DE) or "Drug testing" (DT).

Background of the invention

The need for simulation of biological pathways and chemical reaction pathways by a new approach

Simulation of the kinetics of biological pathways and chemical reaction pathways facilitates the understanding of biological processes, complex chemical reactions, disease processes, and the effect of mutations on biological systems. It can also be used for testing the effect of drugs/molecules on biological systems and disease processes. These are described in "Simulation of prokaryotic genetic circuits", H. H. McAddams, and A. Arkin. Annu. Rev. Biophys. Biomol. Struct. 27, 199-224 (1998).

So far, computational methods appear to be the only practically used approach for simulation of biological pathways as described in "A mathematical model of caspase function in apoptosis", M. Fussenegger, J. E. Bailey, and J. Varner. Nature Biotech. 18, 768-774 (2000); "Mathematical

modeling of epidermal growth factor receptor signaling through the phospholipase C pathway: Mechanistic insights and predictions for molecular interventions", J. M. Haugh, A. Wells, D. A. Lauffenburger. Biotech Bioeng 70, 225-238 (2000); "Computational analysis of biochemical systems, a practical guide for biochemists and molecular biologists", E. O. Voit. Cambridge Press (2000). In these works, the kinetics of binding and reaction events in a pathway, under certain approximations, have also been modeled by a set of first-order linear differential equations.

The kinetics of binding and reaction events in a pathway have also been described by a set of first-order nonlinear ordinary differential equations as proposed in "Systemic properties of ensembles of metabolic networks: application of graphical and statistical methods to simple unbranched pathways", R. Alves, and M. A. Savageau. Bioinformatics 16, 534-547 (2000); and in "Inferring qualitative relations in genetic networks and metabolic pathways", T. Akusu, S. Miyano, and S. Kuhara. Bioinformatics 16, 727-734 (2000).

While they have been successfully used in studying individual signaling and metabolic pathways involving dozens of proteins and other molecules, in the foreseeable future, computational methods are not expected to have the capacity for simulation of large and complex pathways or multiple pathways involving hundreds or more number of proteins as described in "Metabolic networks: a signal-oriented approach to cellular models" J. W. Lengeler. Biol. Chem. 381, 991-920 (2000).

Cellular events frequently involve "crosstalk" interactions among multiple pathways that may include thousands or more proteins and other molecules, and the analysis of such complex biochemical networks is further complicated by the great number of feedback and feedforward loops and of regulatory

networks involved in cellular control, as described in "Metabolic networks: a signal-oriented approach to cellular models" J. W. Lengeler. Biol. Chem. 381, 991-920 (2000) and in "Mathematical modeling of epidermal growth factor receptor signaling through the phospholipase C pathway: Mechanistic insights and predictions for molecular interventions", J. M. Haugh, A. Wells, D. A. Lauffenburger, Biotech Bioeng 70, 225-238 (2000).

The kinetics of such a large network is described by up to thousands of first order nonlinear differential equations. Significantly more numbers of equations may be needed for simulation of cellular and multi-cellular systems.

Given that it is unlikely that the computer resource in the near future is capable of solving the "long-time" kinetics of such a large number of nonlinear equations, a new approach needs to be introduced.

The need for structural optimization and molecular dynamics simulation of biomolecules and nano-molecular-systems by a new approach

Molecular dynamics simulation is a widely used method for simulation and modeling of structures, motions, binding, and thermodynamics of biomolecules and nano-molecular-systems. These are described in "Computer simulation of biomolecular systems: Theretical and experimental applications", Eds. W. F. van Gunsteren, P. K. Weiner, and A. J. Wilkinson. ESCOM Leiden (1993); "Computational nanotechnology", R. C. Merkle, Nanotechnology, 2, 134-141 (1991).; nanostructure engineering as described in "Simulated engineering of nanostructures", D.W. Brenner, S.B. Sinnott, J.A. Harrison, and O.A. Shenderova. Nanotechnology 7 161-167 (1996). It has been used in facilitating the study of a variety of important problems such as protein folding as described in "Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution", Y. Duan and P.A. Kollman. Science 282, 740-744 (1998); drug design as described in

"Binding pathway of retinal to bacterio-opsin: A prediction by molecular dynamics simulations", B. Isralewitz, S. Izrailev, and K. Schulten. Biophys. J. 73, 2972-2979; nanostructure engineering as described in "Simulated engineering of nanostructures", D.W. Brenner, S.B. Sinnott, J.A. Harrison, and O.A. Shenderova. Nanotechnology 7 161-167 (1996).

Nano-molecular-systems refer to: nano-sized or nano-structured materials (such as molecular films, nanotubes, nanoscopic particles with specific electronic, optical or magnetic properties), nano-scale molecular mechanical or manufacturing systems (capable of guiding reactive molecules to 0.1nm precision), other nano-scale devices (molecular motors, carriers, containers, pumps, circuits, tools etc.). These are described in "Nanosystems: molecular machinery, manufacturing, and computation" by K. Eric Drexler (Wiley 1992).

So far, computational methods are the only practically used approach for molecular dynamics simulation as described in "Computer simulation of biomolecular systems: Theretical and experimental applications", Eds. W. F. van Gunsteren, P. K. Weiner, and A. J. Wilkinson. ESCOM Leiden (1993). The molecular dynamics simulation of biomolecules can be described by a set of second-order nonlinear ordinary differential equations as indicated in "A second generation force field for simulation of proteins, nucleic acids, and organic molecules", W. D. Cornell et. al., J. Am. Chem. Soc. 117, 5179-5197 (1995); and in "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations", B. R. Brooks, et.al. J. Comp. Chem. 4, 187-217 (1983). While they have been successfully used in studying some problems of proteins and other molecules, in the foreseeable future, computational methods are not expected to have the capacity for simulation of proteins and nucleic acids for orders of magnitude longer than the microsecond range.

The estimated time scale for folding/unfolding events in proteins, ligand-receptor binding/dissociation, and interaction of nano-molecular-machine with its substrates ranges from hundreds of milliseconds to tens of seconds as described in "Kinetic analysis of the unfolding and refolding of ribonuclease T1 by a stopped-flow double-mixing technique", L. M. Mayr et. al., Biochemistry 35, 5550-5561 (1996); and in "Studies of the binding of actinomycin and related compounds to DNA", W. Muller, and D. M. Crothers, J. Mol. Biol. 35, 251-290 (1968). In contrast, the longest achieved simulation time is 1 microsecond for a small protein, consisting of only 36 amino acids, on a 256-node CRAY T3E supercomputer, as described in "Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution", Y. Duan and P.A. Kollman. Science 282, 740-744 (1998). Hence there is a big gap (5 to 7 orders of magnitude) between simulation time scale and that of certain molecular events. Given that it is unlikely the computer approach in the near future is capable of conducting "long" enough time-scale molecular dynamics simulation, a new approach needs to be introduced.

Summary of the invention

The present invention seeks to address the problems above, and in particular to provide a new approach for the simulation of biological and/or chemical reaction pathway, and for the molecular dynamic simulation of biomolecules and nano-molecular systems by comprising constructing set of binding, reaction, concentration and/or molecular dynamics equations and electronic circuits corresponding to such equations.

According to a first embodiment, the invention provides a method for simulation of at least one biological and/or chemical reaction pathway comprising:

preparing a map of at least one biological and/or chemical reaction pathway;

constructing at least one set of binding and reaction equation from the pathway map;

constructing at least one set of concentration equation for molecules of the pathway map;

constructing an electronic circuit corresponding to every set of equation;

determining simulation of pathway by measuring voltage at two or more connection points of the circuit.

In particular, the binding, reaction and/or concentration equation may be a linear or non-linear first- or second-order ordinary differential equation (ODE).

The electronic circuit may comprise at least one of the following circuit units: linear protein/molecule concentration $\pm kx_i$ unit; protein/molecule concentration power $\pm kx_i^a$ unit; ligand-protein concentration product $\pm Lx_i$ unit; ligand-protein concentration power product $\pm kLx_i^a$ unit; protein-protein concentration product $\pm kx_ix_j$ unit; protein-protein concentration power product $\pm kx_i^ax_j^b$ unit; stochastic rate constant generator unit that replaces a forward/reverse binding/reaction rate constant k by a stochastic value; and

wherein, k is a forward/reverse binding/reaction rate constant, L the concentration of a ligand, x_i and x_j are the concentration of protein/molecule x_i and x_j respectively, a and b are order of power of x_i and x_j.

According to an aspect, the method of the invention further comprises maintaining the voltage level of the circuit between two fixed voltage values. For example, by: multipling scaling factors to the concentration equation;

applying at least one resistor and/or amplifier at one or more connection point of the circuit, thereby scaling-down or scaling–up the voltage of one or more segment of the circuit; and/or applying automatic gain control circuits.

According to another aspect, the method of the invention further comprises adding at least one unit circuit comprising an electronic random noise generator and/or a multiplicator amplifier at one or more connection point of the circuit.

The method of the invention may further comprise determining the effect of at least one drug comprising adding at least one circuit unit associated with a receptor protein at one or more connection point of the circuit.

The method may further comprise determining deficiency, mutation and/or deletion of at least one protein of the biological pathway, comprising setting a voltage ceiling, a voltage range and/or a fixed voltage at one or more connection point of the circuit.

The invention also provides an electronic circuit system for the simulation of at least one biological and/or chemical reaction pathway, comprising at least one electronic circuit representing a set of binding, reaction and/or concentration equation. In particular, an electronic circuit system, wherein the binding, reaction and/or concentration equation is a linear or non-linear first- or second-order ordinary differential equation (ODE).

The electronic circuit may comprise at least one the following circuit units: linear protein/molecule concentration $\pm kx_i$ unit; protein/molecule concentration power $\pm kx_i^a$ unit; ligand-protein concentration product $\pm Lx_i$ unit; ligand-protein concentration power product $\pm kLx_i^a$ unit; protein-protein concentration product $\pm kx_ix_j$ unit; protein-protein concentration power product $\pm kx_i^ax_j^b$ unit; stochastic

rate constant generator unit that replaces a forward/reverse binding/reaction rate constant k by a stochastic value; and

wherein, k is a forward/reverse binding/reaction rate constant, L the concentration of a ligand, x_i and x_j are the concentration of protein/molecule x_i and x_j respectively, a and b are order of power of x_i and x_j respectively.

According to another embodiment, the invention provides a method for molecular dynamics simulation of biomolecules and/or nano-molecular systems comprising:

constructing at least one set of equation representing the molecular dynamics of at least one molecule of the biomolecules and/or the nano-molecular systems;

constructing an electronic circuit representing every set of equation;

determining molecular dynamics simulation by measuring voltage at two or more connection points of the circuit.

The equation may be a linear or non-linear second order ordinary differential equation (ODE).

According to an aspect, in the method of the invention, the electronic circuit comprises at least one atom-position circuit unit, wherein the atom-position circuit unit represents the position of an atom of a molecule or a molecular system.

In particular, the atom-position circuit unit comprises at least one atom-atom interaction circuit subunit, the atom-atom interaction circuit subunit representing a sub-unit of atom-atom interactions within a molecule or a molecular system and comprising internal bond stretch, angle bending, torsion, non-bonded units; and/or bond stretch, angle bending, and/or torsion units; between at least two nearest sub-units of a molecule.

Further, the atom-atom interaction circuit subunit may represent a term in the molecular dynamics equation, and wherein the atom-atom interaction circuit subunit comprises at least one of the following: bond stretch x unit, bond stretch y unit, bond stretch z unit, angle bending x type-A unit, angle bending x type-B unit, angle bending y type-A unit, angle bending y type-B unit, angle bending z type-A unit, angle bending z type-B unit, torsion x type-A unit, torsion x type-B unit, torsion y type-A unit, torsion y type-B unit, torsion z type-A unit, torsion z type-B unit, non-bonded x unit, non-bonded y unit, non-bonded z unit, hydrogen-bond x unit, hydrogen-bond y unit, and hydrogen-bond z unit; and wherein x, y, and z represent the coordinates of each atom of the molecule, and type-A represents the case of the atom being in the middle-position of an angle bending or torsion connection with other atoms, and type-B represents the case of the atom being in the end-position of an angle bending or torsion connection with other atoms.

The method may further comprise maintaining the voltage level in the circuit between two fixed voltage values. For example, by wherein x, y and z represent the coordinate of the molecule, and the voltage level of the circuit is maintained between two fixed voltage values by:

applying scaling factors to the x, y and z coordinates and to the molecular dynamic equation;

applying at least one resistor and/or amplifier at one or more connection point of the circuit, thereby scaling-down or –up the voltage of one or more segment of the circuit; and/or

applying automatic gain control circuits.

The invention also provides a circuit group representing the interaction pattern in the chemical structure of a molecule or a sub-unit of interaction pattern in the chemical structure of a molecule comprising:

a bond stretch connection between each atom pair of the molecule covalently bonded to each other;

an angle bending connection pair between a first atom and other two atoms;

a torsion connection bundle between a first atom and other three atoms; and

a non-bonded connection between each atom pair whose atoms are at least four bonds away from each other.

The electronic circuit may comprise at least one circuit unit, the circuit unit comprising at least one circuit group, the circuit group representing a sub-unit of interaction pattern in the chemical structure of a molecule and comprising internal bond stretch, angle bending, torsion, non-bonded units; and/or bond stretch, angle bending, and/or and torsion units; between at least two nearest sub-unit of a molecule.

The circuit unit may represent a term in the molecular dynamic equation, and wherein the circuit unit comprises at least one of the following: bond stretch x unit, bond stretch y unit, bond stretch z unit, angle bending x type-A unit, angle bending x type-B unit, angle bending y type-A unit, angle bending y type-B unit, angle bending z type-A unit, angle bending z type-B unit, torsion x type-A unit, torsion x type-B unit, torsion y type-A unit, torsion y type-B unit, torsion z type-A unit, torsion z type-B unit, non-bonded x unit, non-bonded y unit, non-bonded z unit, hydrogen-bond x unit, hydrogen-bond y unit, and hydrogen-bond z unit; and wherein x, y, and z represent the coordinates of each atom of the molecule, and type-A unit represents the case of the atom being in the middle-position of an angle bending or torsion connection with other atoms, and type-B represents the case of the atom being in the end-position of an angle bending or torsion connection with other atoms.

The invention further relates to a method for the manufacture of an electronic circuit representing at least one biomolecule and/or nano-molecular system, the electronic circuit comprising at least a unit circuit comprising at least a circuit group, wherein the circuit group represents a sub-unit of interaction pattern in the chemical structure of a molecule or a molecular system comprising:

introducing a bond stretch between each atom of a pair of atoms covalently bonded to each other;

introducing an angle bending connecting pair between a first atom and other two atoms;

introducing a torsion connection bundle between a first atom and other three atoms; and

introducing a non-bonded connection between each atom pair whose atoms are at least 4 bonds away from each other.

Brief description of the drawings

Figure 1 is a flow diagram that illustrates one embodiment of a process for simulating a single or multiple biological or chemical reaction pathways by electronic circuits (100).

Figure 2 is the map of an illustrative biological pathway, part of the caspase pathway.

Figure 3A, 3B are the maps of the actual circuits for an illustrative biological pathway, part of the caspase pathway, wherein Figure 3A represents the circuit for Equation I ($\frac{dx_1}{dt} = k_{1-}x_2 - k_{1+}Lx_1 + \omega_1$) and Figure 3B represents the circuit for Equation II ($\frac{dx_2}{dt} = k_{1+}Lx_1 - (k_{1-} + k_2)x_2$).

Figure 4 is a map representing an electronic emulator for the entire pathway governed by twenty ODEs, of the caspase pathway, where blocks Equation 1 to 20 represent the circuits that emulate the twenty ODEs.

Figure 5 is the measured voltage changes, representing variations of protein concentrations in the illustrative biological pathway, from the designed circuit for this pathway.

Figure 6 is a schematic diagram of the construction of a section of electronic circuits representing an ODE associated with a binding/reaction event in a biological or chemical pathway. Note that switches I1 and I2 are closed at t=0 when the initial condition is $x_1=\alpha$ or $dx_i/dt=\beta$. Each switch can also be used when x_i or dx_i/dt is set at a fixed value or a fixed range.

Figure 7 is a schematic diagram of the circuit for the first ODE of the illustrative caspase pathway. Note that switches I1 and I2 are closed at t=0 when the initial condition is $x_1=\alpha$ or $dx_i/dt=\beta$. Each switch can also be used when x_i or dx_i/dt is set at a fixed value or a fixed range.

Figure 8 is a schematic diagram of the design plans for the special circuit units representing each term in the ODE of a binding/reaction event in a pathway.

Figure 9 is a schematic diagram of the design plans for the special circuit units representing each term in the ODE of a binding/reaction event in a pathway. The random number ξ is in the range: $0<\xi<1$.

Figure 10 is a schematic diagram of the construction of a section of electronic circuits representing a molecular dynamics simulation ODE for the x-coordinate of an atom.

Figure 11 is a schematic diagram of the construction of a section of electronic circuits representing a molecular dynamics simulation ODE for the y-coordinate of an atom.

Figure 12 is a schematic diagram of the construction of a section of electronic circuits representing a molecular dynamics simulation ODE for the z-coordinate of an atom.

Figure 13 is a schematic diagram of the bond stretch connections of the circuit group representing a subunit, amino acid, of an illustrative biomolecule, protein.

Figure 14 is a schematic diagram of the angle bending connections of the circuit group representing a subunit, amino acid, of an illustrative biomolecule, protein.

Figure 15 is a schematic diagram of the torsion connections of the circuit group representing a subunit, amino acid, of an illustrative biomolecule, protein.

Figure 16 is the structure of an illustrative nano-molecular-machine, a rudimentary mimic of the enzyme ribonuclease, which has been constructed from the bucket-shaped cyclodextrin molecule (a). midazole functional groups on the rim (X in (b)) help to catalyze the region specific hydrolysis of a cyclic phosphodiester (b), which is bound in the hydrophobic cavity.

Figure 17 is a schematic diagram of the bond stretch connections of the circuit group representing a subunit of an illustrative nano-molecular-machine, a rudimentary mimic of the enzyme ribonuclease.

Figure 18 is a schematic diagram of the angle-bending connections of the circuit group representing a subunit of an illustrative nano-molecular-machine, a rudimentary mimic of the enzyme ribonuclease.

Figure 19 is a schematic diagram of the torsion connections of the circuit group representing a subunit of an illustrative nano-molecular-machine, a rudimentary mimic of the enzyme ribonuclease.

Figure 20 is a schematic diagram of the bond stretch x unit, representing the x-component of a bond stretch term.

Figure 21 is a schematic diagram of the bond stretch y unit, representing the y-component of a bond stretch term.

Figure 22 is a schematic diagram of the bond stretch z unit, representing the z-component of a bond stretch term.

Figure 23 is a schematic diagram of the angle bending x type-A unit, representing the x-component of a bond angle bending term in which the atom is in the middle.

Figure 24 is a schematic diagram of the angle bending x type-B unit, representing the x-component of a bond angle bending term in which the atom is at one end.

Figure 25 is a schematic diagram of the angle bending y type-A unit, representing the y-component of a bond angle bending term in which the atom is in the middle.

Figure 26 is a schematic diagram of the angle bending y type-B unit, representing the y-component of a bond angle bending term in which the atom is at one end.

Figure 27 is a schematic diagram of the angle bending z type-A unit, representing the z-component of a bond angle bending term in which the atom is in the middle.

Figure 28 is a schematic diagram of the angle bending z type-B unit, representing the y-component of a bond angle bending term in which the atom is at one end.

Figure 29 is a schematic diagram of the torsion x type-A unit, representing the x-component of a bond torsion term in which the atom is in the middle.

Figure 30 is a schematic diagram of the torsion x type-B unit, representing the x-component of a bond torsion term in which the atom is at one end.

Figure 31 is a schematic diagram of the torsion y type-A unit, representing the y-component of a bond torsion term in which the atom is in the middle.

Figure 32 is a schematic diagram of the torsion y type-B unit, representing the y-component of a bond torsion term in which the atom is at one end.

Figure 33 is a schematic diagram of the torsion z type-A unit, representing the z-component of a bond torsion term in which the atom is in the middle.

Figure 34 is a schematic diagram of the torsion z type-B unit, representing the z-component of a bond torsion term in which the atom is at one end.

Figure 35 is a schematic diagram of the non-bonded x unit, representing the x-component of a non-bonded interaction term.

Figure 36 is a schematic diagram of the non-bonded y unit, representing the y-component of a non-bonded interaction term.

Figure 37 is a schematic diagram of the non-bonded z unit, representing the z-component of a non-bonded interaction term.

Figure 38 is a schematic diagram of the H-bond x unit, representing the x-component of a hydrogen-bond term.

Figure 39 is a schematic diagram of the H-bond y unit, representing the y-component of a hydrogen-bond term.

Figure 40 is a schematic diagram of the H-bond z unit, representing the z-component of a hydrogen-bond term.

Figure 41 is a schematic diagram of the rABC circuit unit and the AqA circuit unit.

Figure 42 is a schematic diagram of the Ua circuit unit, ss circuit unit, and si circuit unit.

Figure 43 is a schematic diagram of the pq circuit unit.

Figure 44 is a schematic diagram of the ϕ circuit unit and the $\sin\phi$ circuit unit.

Detailed description of the invention

Method for simulation of biological and/or chemical reaction pathway

According to a first embodiment, the present invention provides a method for simulation of a single or multiple biological and/or chemical reaction pathways.

It has particular application in simulation of large networks of pathways that are beyond the scope of computer methods.

It provides a potentially useful means for facilitating the study of cellular and multi-cellular processes, disease processes, effect of gene or protein mutations on biological systems, and effect of drugs on biological systems.

A single or a set of biological pathways and/or chemical reaction pathways can be represented by a map composed of a network of boxes or labels (reaction objects) connected by arrows (binding/reaction paths). This kind of representation is widely used in popular biological pathway databases such as SPAD (http://www.grt.kyushu-u.ac.jp/spad/) and KEGG (http://www.genome.ad.jp/kegg/kegg2.html) and as described in "Mathematical modeling of epidermal growth factor receptor signaling through the phospholipase C pathway: Mechanistic insights and predictions for molecular interventions", J. M. Haugh, A. Wells, D. A. Lauffenburger. Biotech Bioeng 70, 225-238 (2000). Each binding/reaction event in these pathways can be described by a binding/reaction equation, whose form depends on whether the event is reversible or irreversible.

The set of binding and/or reaction (binding/reaction) equations for pathways can be modeled by a set of linear and/or non-linear first-order and/or second-

order ordinary differential equations. For example, non-linear first-order ordinary differential equations (ODEs) derived from the balance equation and generalized mass-action kinetics as described in "Understanding the control of metabolism" D. Fell, Portland Press, London, UK (1996) and in "The regulation of cellular systems" R. Heinrich, and S Schuster, Chapman & Hall, New York (1996). Each ODE describes the time-dependent behaviour of the concentration of a protein/molecule in the pathways. Further, equations representing the concentration of the molecules of the pathway map of the biological and/or chemical reaction pathway are also provided. The concentration equations can also be at least a set of linear and/or non-linear first-order and/or second-order ordinary differential equations. For example, non-linear first-order ordinary differential equations (ODEs).

According to this embodiment of the invention, at least one set of electronic circuits represents the at least one set of the above equations. In particular, binding, reaction and/or concentration non-linear ODEs.

Electronic circuits have been proposed to solve certain first-order and second-order nonlinear ordinary differential equation as described in "The transition from solitons to chaos in the solution of the logistic equation" M. I. Sobhy, and A. S. Burman. Int. J. Bifurcation Chaos 10, 2823-2829 (2000) as well as linear ordinary differential equation (up to second-order) as described in "The analog computer as an aid to the teaching elementary quantum mechanics", M. K. Summers. Phys. Educ 13, 22-27 (1996).

However, electronic circuits have never been disclosed nor suggested for the purpose of simulation of biological and/or chemical reaction pathway.

Operational amplifiers, function circuits and other electronic components can be used to construct operations of differentiation, integration, sum,

subtraction, multiplication, inversion, exponential, logarithm, power, and others as described in "Electronic circuits and design" D. A. Neamen. McGraw Hill (2001), "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGraw-Hill (1976), "Modern operational circuit design" J.I. Smith, Wiley-Interseicne (1971). Hence, a set of nonlinear or linear ordinary differential equations composed of these terms can be represented by a set of electronic circuits composed of these operational amplifiers and other electronic components. The time-dependent behaviour of the concentration of a protein/molecule can be measured by the voltage at a specific point in the circuits.

Accordingly, for simulation of biological and/or chemical reaction pathways, the time-dependent behavior of the concentration of one or more reaction objects, defined as proteins/nucleic acids/ligands/substrates/reactants/products, in the pathways can be derived by measuring the voltage at specific points in the circuits.

In order to accurately describe the microscopic kinetics of binding or reaction events in pathways, all possible states for each protein or molecule need to be considered. Examples of these states include active, inactive, ligand/substrate-bound, dimmer, and any other state. The corresponding circuit section for the concentration in a particular state is typically composed of up to a few dozen operational amplifiers, a similar number of connection points (two for measuring its voltage, a few for connection to external voltage sources, and several for connection to and from other circuit units), and a number of small components such as switches etc. The circuits for a large pathway or multiple pathways may thus contain up to thousands or more such sections connected to each other and to the external voltage sources, which is feasible to construct.

Accordingly, the first embodiment provides a method for simulation of at least one biological and/or chemical reaction pathway comprising:

preparing a map of at least one biological and/or chemical reaction pathway;

constructing at least one set of binding and reaction equation from the pathway map;

constructing at least one set of concentration equation for molecules of the pathway map;

constructing an electronic circuit corresponding to every set of equation;

determining simulation of pathway by measuring voltage at two or more connection points of the circuit.

The binding, reaction and/or concentration equation may be a linear or non-linear first- or second-order ordinary differential equation (ODE). Preferably, the binding, reaction and/or concentration equation is a non-linear first-order ordinary differential equation (ODE).

In a further aspect of this invention, methods are disclosed for maintaining the voltages in the circuits (concentrations of molecules and functions of concentrations) within the allowed range. Scaling factors (which are non-zero real numbers) are applied to the concentrations and to the kinetic equations so as to keep these concentrations and each term in the equations within the allowed range. In addition, resistors and amplifiers can be used to scale-down and scale-up the voltages at certain segment of the circuits. Moreover automatic gain control circuits can also be used to ensure the voltages are within the required range. The automatic gain control circuits are described in "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGram-Hill, 1976.

Accordingly, the invention provides a method further comprising maintaining the voltage level of the circuit between two fixed voltage values.

In particular, the voltage level of the circuit is maintained between two fixed voltage values by:

multipling scaling factors to the concentration equation;

applying at least one resistor and/or amplifier at one or more connection point of the circuit, thereby scaling-down or scaling–up the voltage of one or more segment of the circuit; and/or

applying automatic gain control circuits.

In a further aspect of this invention, stochastic effect on the binding/reaction events, which is important in cases involving low concentration or short-time event, can be taken into account by replacing the relevant binding/reaction rate constants by stochastic values (random variations around average rate constants determined by experiments or by theoretical methods). This can be accomplished in the circuits by adding at specific sites special circuit units composed of an electronic random noise generator device, such as Wayne's random noise generators, and a multiplication amplifier.

Accordingly, the method of the invention further comprises adding at least one unit circuit comprising an electronic random noise generator and/or a multiplicator amplifier at one or more connection point of the circuit.

In a further aspect of the invention, design plans for several circuit units are introduced to represent one or more, or even all possible terms in a circuit equation. Accordingly, in the method of the invention, the electronic circuit may comprise at least one of the following circuit units: linear protein/molecule concentration $\pm kx_i$ unit; protein/molecule concentration power $\pm kx_i^a$ unit;

ligand-protein concentration product $\pm Lx_i$ unit; ligand-protein concentration power product $\pm kLx_i^a$ unit; protein-protein concentration product $\pm kx_ix_j$ unit; protein-protein concentration power product $\pm kx_i^ax_j^b$ unit; stochastic rate constant generator unit that replaces a forward/reverse binding/reaction rate constant k by a stochastic value; and

wherein, k is a forward/reverse binding/reaction rate constant, L the concentration of a ligand, x_i and x_j are the concentration of protein/molecule x_i and x_j respectively, a and b are order of power of x_i and x_j.

By introducing additional circuit sections or by modifying the circuits, it is possible to simulate events such as the effect of drugs, protein or gene mutations, and diseases on a pathway or a group of pathways.

Therefore, in a further aspect of this invention, effect of drugs can be simulated by adding one or more circuit units (the design plan for each of these units is disclosed) at specific sites in the circuits that are associated with the respective receptor proteins. These added units represent the forward and reverse binding/reaction process between these drugs and their respective receptor proteins.

Accordingly, the method of the invention, further comprises determining the effect of at least one drug comprising adding at least one circuit unit associated with a receptor protein at one or more connection point of the circuit.

In a further aspect of this invention, deficiency or mutation of one or more proteins can be simulated by setting a voltage ceiling, or a voltage range, or a fixed voltage value at specific points in the circuits.

Accordingly, the method of the invention further comprises determining deficiency, mutation and/or deletion of at least one protein of the biological pathway, comprising setting a voltage ceiling, a voltage range and/or a fixed voltage at one or more connection point of the circuit.

The method of claim 1, wherein the biological pathway comprises at least one object, and wherein the object is protein, nucleic acid, ligand, substrate, inhibitor or antagonist, activator or agonist, reactant and/or reaction product.

According to the first embodiment of the invention, it is also provided an electronic circuit system for the simulation of at least one biological and/or chemical reaction pathway, comprising at least one electronic circuit representing a set of binding, reaction and/or concentration equation.

In the electronic circuit system, the binding, reaction and/or concentration equation may be a linear or non-linear first- or second-order ordinary differential equation (ODE). Preferably, a non-linear first-order ODE.

According to a particular aspect, the invention relates to the disclosure of design plans for the construction of circuit units that represent one or more or even all possible terms in the ODEs describing a biological or a chemical reaction pathway.

Accordingly, in the electronic circuit system of the invention, the electronic circuit may comprise at least one, or even all, the following circuit units: linear protein/molecule concentration $\pm kx_i$ unit; protein/molecule concentration power $\pm kx_i^a$ unit; ligand-protein concentration product $\pm Lx_i$ unit; ligand-protein concentration power product $\pm kLx_i^a$ unit; protein-protein concentration product $\pm kx_ix_j$ unit; protein-protein concentration power product $\pm kx_i^ax_j^b$ unit; stochastic

rate constant generator unit that replaces a forward/reverse binding/reaction rate constant k by a stochastic value; and

wherein, k is a forward/reverse binding/reaction rate constant, L the concentration of a ligand, x_i and x_j are the concentration of protein/molecule x_i and x_j respectively, a and b are order of power of x_i and x_j respectively.

Referring now to the drawings and, in particular, to Figure 1, there is disclosed a process 100 for carrying out a method for the simulation of a single or multiple biological or chemical reaction pathways. In the process 100, simulation of the pathways is conducted on a specifically designed set of electronic circuits by the procedure as described below and given in the flow chart of Figure 1.

The process 100 begins at a state 110, wherein the pathways are presented as a map of a network of boxes or labels (proteins or molecules) connected by arrows (binding/reaction paths). Figure 2 shows the map of a part of caspase pathway as an illustration.

The process 100 then moves to a sate 120 wherein the binding or reaction equation for every process in the pathway is constructed from the pathway map. Based on whether it is reversible or irreversible, the equation of an individual binding /reaction step takes the form:

$$\text{Reversible binding/reaction: } R_1 + R_2 + \ldots \underset{K_-}{\overset{K_+}{\rightleftharpoons}} P_1 + P_2 + \ldots$$

$$\text{Irreversible binding/reaction: } R_1 + R_2 + \ldots \xrightarrow{K_+} P_1 + P_2 + \ldots$$

where R_1, R_2, ... are binding/reaction objects (proteins/molecules etc) and P_1, P_2, ... are product objects (proteins/molecules etc), K_+ and K_- are forward and reverse binding/reaction rate constant respectively.

The first few reaction equations for the illustrative caspase pathway in Figure 2 are given as:

Event 1: $X_1 + L \underset{k_{1-}}{\overset{K_{1+}}{\rightleftharpoons}} X_2$

Event 2: $X_2 \overset{K_{2+}}{\dashrightarrow} X_3$

Event 3: $X_3 + X_4 \underset{K_{3-}}{\overset{K_{3+}}{\rightleftharpoons}} X_5$

Event 4: $X_4 + X_5 \underset{K_{4-}}{\overset{K_{4+}}{\rightleftharpoons}} X_6$

where L = Fas ligand, x_1=ligand-free Fas surface receptor, x_2=ligand-bound Fas surface receptor, x_3=clustered ligand-bound Fas surface receptor, x_4=FADD protein, x_5=complex of FADD-Fas receptor, ..., and k_{i+} and k_{i-} are the forward and reverse reaction rate constant in the i-th path respectively. These rate constants can either be obtained from experimental data, or from theoretical computations, or estimated by empirical means based on statistical analysis of available experimental data of available of relevant reactions.

In general, the reaction equation for the i-th path is:

$$\textbf{Path i:}\quad \Sigma_j\, a_{ij}\, x_j \underset{k_{i-}}{\overset{k_{i+}}{\rightleftharpoons}} \Sigma_k\, b_{ik}\, x_k$$

where a_{ij} and b_{ik} are the stoichiometric coefficient of the j-th and k-th reaction objects respectively; j=1, ..., N_j; k=1, ...,N_k; and N_j and N_k are number of reaction objects and number of product objects respectively.

The process 100 then moves to a state 130 where in the corresponding kinetic equations for the concentration of each protein or molecule are derived from the balance equation and generalized mass-action kinetics as described in "Understanding the control of metabolism" D. Fell, Portland Press, London, UK (1996) and in "The regulation of cellular systems" R. Heinrich, and S Schuster, Chapman & Hall, New York (1996). Each equation is a non-linear

first order ordinary differential equation (ODE) of the concentration of a protein/molecule in the pathways. In general, for a set of equations involving the n-th reaction object:

$$a_{in}\, X_n + \Sigma_j\, a_{ij}\, X_j \underset{K_{i-}}{\overset{K_{i+}}{\rightleftharpoons}} \Sigma_k\, b_{ik}\, x_k$$

$$\Sigma_{j'}\, a_{ij'}\, X_{j'} \underset{K_{i'-}}{\overset{K_{i'+}}{\rightleftharpoons}} a_{i'n}\, X_n + \Sigma_{k'}\, b_{i'k'}\, x_{k'}$$

the corresponding ODE is:

$$dx_n/dt = k_{i-} \Pi_k (x_k)^{bik} + k_{i'+} \Pi_{j'} (x_{j'})^{aij'} + \ldots - k_{i+} (x_n)^{ain} \Pi_j (x_j)^{aij} - k_{i'-} (x_n)^{ai'n} \Pi_{k'} (x_{k'})^{bi'k'} - \ldots + \omega_n$$

where x_n is the concentration of object x_n, ω_n is its rate of expression/production, a_{ij} and b_{ij} are the power of the concentration in cases where the reaction rate is determined by power law representation.

The first few ODEs for the illustrative caspase pathway are:

$$dx_1/dt = k_{1-} X_2 - k_{1+} L x_1 + \omega_1 \qquad dx_3/dt = k_2 X_2 + k_{3-} X_5 - k_{3+} x_3 x_4$$
$$dx_2/dt = k_{1+} L X_1 - (k_{1-} + k_2) x_2 \qquad dx_4/dt = k_{3-} X_5 + k_{5-} X_6 - k_{3+} x_3 x_4 - k_{5+} x_4 x_5 + \omega_4$$

where the ligand concentration L and rate of expression ω_n of protein x_n are assumed to be time-dependent function.

Referring back to Figure 1, the process 100 then moves to a state 140 wherein a network of electronic circuits are constructed to represent the ODEs of a pathway or a network of pathways. This circuit network is composed of specially designed circuit units as disclosed in this patent and known components such as operational amplifiers, voltage sources representing time-dependent distribution of ligands or molecules, and other electronic

circuit components such as switches and voltage sources. Each equation has a circuit, which is connected to other circuits. An equation normally contains no more than a dozen terms or variables, and thus its circuit needs to have no more than a dozen connection points to other circuits. Each circuit starts with a point x_I, which is connected to a switch that controls the initial condition, or the range, or the fixed value of voltage. This point is connected to other parts of the circuit by a combination of one or more of the disclosed specially designed circuit units together with a differentiator and a summing amplifier as illustrated in Figure 6. Figure 7 shows the actual circuit design for the first ODE of the illustrative caspase pathway. The differentiator, summing amplifier, and inverting amplifier are described in "Electronic circuit analysis and design, 2nd edition" D.A. Neamen, McGraw-Hill, 2001. The multiplier is described in "Fundamentals of linear circuits" T.L. Floyd, Macmillan Publishing Co. 1992.

The process 100 then moves to a simulation state 150 wherein the time-dependent kinetics of the concentration of one or more proteins or molecules in a pathway or a group of pathways can be determined by measurement of voltages at specific sites of the circuits. For the circuit illustrated in Figure 6, the voltage at x_i corresponds to the concentration of protein/molecule x_i. For the circuit illustrated in Figure 7, the voltage at x_1 corresponds to the concentration of protein x_1.

Construction Plan for the Specially Designed Circuit Units Representing Different Terms in an ODE For Pathway Simulation

Referring now to Figure 8 and Figure 9, there are disclosed design plans for construction of the special circuit units according to the invention representing all possible terms of an ordinary differential equation associated with a

biological and/or chemical reaction pathway. The operational amplifiers and random noise generator in these units are all described in the literature. The inverting amplifier and noninverting amplifier are described in "Electronic circuit analysis and design, 2nd edition" D.A. Neamen, McGraw-Hill, 2001. The multiplier is described in "Fundamentals of linear circuits" T.L. Floyd, Macmillan Publishing Co. 1992. The exponentiator is described in "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGram-Hill, 1976. The random noise generator is described in "A Noise-Based IC Random Number Generator for Applications in Cryptography," Petrie, C.S. and Connelly, J.A., IEEE Transactions on Circuits and Systems, Part I, Vol. 47, May 2000, pp. 615-621.

Methods for Maintaining the Voltages in the Circuits Within Allowed Range

Normally the concentration of a molecule can be scaled into the allowed voltage range of a circuit. In case that this may not be so, or in case that the derivative of concentration exceeds the voltage range, or in case that some terms in the kinetic equation may exceed the voltage range, the following methods can be used to maintain the voltage level of the circuit between two voltage values:

(1) Applying scaling factors to the concentrations or to the pathway simulation ODEs so as to keep the concentrations and each term in the equations within the allowed range.

This can be accomplished by multiplying an appropriately selected small number δ to the pathway simulation ODEs (that is, using non-zero real numbers)

(2) Resistors are used to scale-down the voltages at the input end of a segment of the circuits and amplifiers are used at the output end of the segment to sale-up the voltages to the proper level.

This can be accomplished by applying at least one resistor and/or amplifier at one or more connection point of the circuit, thereby scaling-down or scaling-up the voltage of one or more segment of the circuit.

(3) Automatic gain control circuits can be added to the circuits to ensure the voltages are within the required range.

Simulation of Effect of Drugs

The effect of drugs on a pathway or a group of pathways can be simulated by adding ligand-protein concentration product units to the circuit sections of the drug receptor proteins. This corresponds to case 3 in Figure 6. The ligand-protein concentration product units are shown in Figure 9.

Illustrative example: Electronic pathway emulator of a caspase pathway

As mentioned before in the description of the first embodiment of the invention, a pathway can be described by a set of binding/reaction equations. This set of binding and/or reaction equations can be further modeled by a set of ODEs. An illustrative biological pathway is part of caspase pathway in Figure 2 is governed by the following 20 simultaneous ODEs.

(I) $$\frac{dx_1}{dt} = k_{1-}x_2 - k_{1+}Lx_1 + \omega_1$$

(II) $$\frac{dx_2}{dt} = k_{1+}Lx_1 - (k_{1-} + k_2)x_2$$

(III) $$\frac{dx_3}{dt} = k_2x_2 + k_{3-}x_5 - k_{3+}x_3x_4$$

(IV) $$\frac{dx_4}{dt} = k_{3-}x_5 + k_{5-}x_6 - k_{3+}x_3x_4 - k_{5+}x_4x_5 + \omega_4$$

(V) $$\frac{dx_5}{dt} = k_{5-}x_6 + k_{3+}x_3x_4 - k_{3-}x_5 - k_{5+}x_4x_5$$

(VI) $$\frac{dx_6}{dt} = k_{6-}x_8 + k_9x_9 + k_{10-}x_{11} + k_{5+}x_4x_5 - k_{5-}x_6 - k_{6+}x_6x_7 - k_{10+}x_6x_{10}$$

(VII) $$\frac{dx_7}{dt} = k_{6-}x_8 + k_{8-}x_9 - k_{6+}x_6x_7 - k_{8+}x_7x_8 + \omega_7$$

(VIII) $$\frac{dx_8}{dt} = k_{8+}x_9 + k_{11-}x_{12} + k_{6+}x_6x_7 - k_{6-}x_8 - k_{8-}x_7x_8 - k_{11+}x_8x_{10}$$

(IX) $$\frac{dx_9}{dt} = k_{8-}x_7x_8 - (k_{8+} + k_9)x_9$$

(X) $$\frac{dx_{10}}{dt} = k_{10-}x_{11} + k_{11-}x_{12} - k_{10+}x_6x_{10} - k_{11+}x_8x_{10} + \omega_{10}$$

(XI) $$\frac{dx_{11}}{dt} = k_{10+}x_6x_{10} - k_{10-}x_{11}$$

(XII) $$\frac{dx_{12}}{dt} = k_{11+}x_8x_{10} - k_{11-}x_{12}$$

(XIII) $$\frac{dx_{13}}{dt} = k_9x_9 + k_{13-}x_{15} + k_{15}x_{15} + k_{17-}x_{18} - k_{13+}x_{13}x_{14} - k_{17+}x_{13}x_{17}$$

(XIV) $$\frac{dx_{14}}{dt} = k_{13-}x_{15} - k_{13+}x_{13}x_{14} + \omega_{14}$$

(XV) $$\frac{dx_{15}}{dt} = k_{13+}x_{13}x_{14} - (k_{13-} + k_{15})x_{15}$$

(XVI) $$\frac{dx_{16}}{dt} = k_{15}x_{15} + k_{19-}x_{20} - k_{19+}x_{16}x_{19}$$

(XVII) $$\frac{dx_{17}}{dt} = k_{17-}x_{18} - k_{17+}x_{13}x_{17} + \omega_{17}$$

(XVIII) $$\frac{dx_{18}}{dt} = k_{17+}x_{13}x_{17} - k_{17-}x_{18}$$

(XIX) $$\frac{dx_{19}}{dt} = k_{19-}x_{20} - k_{19+}x_{16}x_{19} + \omega_{19}$$

(XX) $$\frac{dx_{20}}{dt} = k_{19+}x_{16}x_{19} - k_{19-}x_{20}$$

where the ligand concentration L and rate of expression ω_n can change with time (a time-dependent function). Variable x1 to x20 represent the protein concentration at the respective paths. This set of nonlinear ODEs can be represented by a corresponding set of electronic circuits composed of different function blocks. Thus, an electronic biological pathways emulator can be built to mimic and study the biological process.

Among the given 20 simultaneous ODEs, Eqn. (I) to (IV) are of typical types. These typical equations can be emulated using the circuits shown in Figures 3A, 3B and in Figure 4.

The emulator circuit has been constructed and the measured voltages, representing variations of protein concentrations in the illustrative biological pathway, are shown in Figure 5, where voltage x1, x2, ... x20 represent the concentration of each protein in a particular state (active or inactive). The results have shown the expected behaviour as found in computer simulation.

Molecular dynamics simulations of biomolecules and/or nano-molecular-system

According to a second embodiment, the invention relates to molecular dynamics simulations of biomolecules and/or nano-molecular-systems.

Nano-molecular-systems refer to: nano-sized or nano-structured materials (such as molecular films, nanotubes, nanoscopic particles with specific

electronic, optical or magnetic properties), nano-scale molecular mechanical or manufacturing systems (capable of guiding reactive molecules to 0.1nm precision), other nano-scale devices (molecular motors, carriers, containers, pumps, circuits, tools etc.). These are described in "Nanosystems: molecular machinery, manufacturing, and computation" by K. Eric Drexler (Wiley 1992).

The molecular dynamics of a protein/nucleic acid/organic molecule can be represented by a set of linear and/or non-linear second-order ordinary differential equations (ODEs) derived from the Newton's second law of motion given the energy functions, force fields and a starting 3D structure of the system, as described in "A second generation force field for simulation of proteins, nucleic acids, and organic molecules", W. D. Cornell et. al., J. Am. Chem. Soc. 117, 5179-5197 (1995); and in "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations", B. R. Brooks, et.al. J. Comp. Chem. 4, 187-217 (1983). Each ODE describes the time-dependent change of the x- or y- or z- coordinate of each atom in a studied molecule.

According to the second embodiment of the invention, a corresponding set of electronic circuits is proposed to represent this set of linear or non-linear ODEs. This set of electronic circuits is composed of operational amplifiers, function circuits and other electronic circuit elements. A combination of one or more specially designed circuit units (composed of operational amplifiers, function circuits and other electronic components together with differentiator and summing amplifier) is used to represent each term of the ODEs. The time-dependent position of each atom in a protein/nucleic acid/organic molecule can be measured by the voltage at a specific point in the circuits.

Electronic circuits have been proposed to solve certain first-order and second-order nonlinear ordinary differential equation as described in "The transition

from solitons to chaos in the solution of the logistic equation" M. I. Sobhy, and A. S. Burman. Int. J. Bifurcation Chaos 10, 2823-2829 (2000) as well as linear ordinary differential equation (up to second-order) as described in "The analog computer as an aid to the teaching elementary quantum mechanics", M. K. Summers. Phys. Educ 13, 22-27 (1996). However, electronic circuits have never been disclosed nor suggested for the purpose of molecular dynamics simulations of biomolecules and/or nano-molecular systems.

If a 3D structure is not available, a starting structure of a molecule can be generated from its molecular bond connectivity profile and from the standard average bond length, angle and torsion. Solution of these molecular dynamics simulation equations gives the time evolution of the position of each atom of a studied molecule. All the terms in this equation can be represented by a combination of existing operational amplifiers, function circuits and other electronic components. Hence it is feasible to use electronic circuits to solve these equations and thus conduct molecular dynamics simulations.

A protein normally consists of thousands of atoms. Each atom has three equations describing its x-, y-, or z- coordinate respectively. The corresponding circuit section for each coordinate is typically composed of up to a few dozen bond-stretch/angle/torsion units plus up to thousands of non-bonded units. If a force field with separate hydrogen bond terms is used, then this circuit section also contains a few additional units for hydrogen bonds. If explicit solvent is used, then this circuit section also contains additional non-bonded units related to the non-bonded interactions with water molecules. Each unit consists of approximately a few dozen operational amplifiers and function amplifiers, a similar number of connection points (two for measuring its voltage, a few for connection to external voltage sources, and the rest for connection to and from other circuit units), and a number of small components such as switches etc. The circuits for a typical protein may thus contain up to

thousands or more such sections connected to each other and to the external voltage sources, which is feasible to construct.

Operational amplifiers, function circuits and other electronic components can be used to construct operations of differentiation, integration, sum, subtraction, multiplication, inversion, exponential, logarithm, power, and others as described in "Electronic circuits and design" D. A. Neamen. McGraw Hill (2001), "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGraw-Hill (1976), "Modern operational circuit design" J.I. Smith, Wiley-Interseicne (1971). Hence, a set of nonlinear or linear ordinary differential equations composed of these terms can be represented by a set of electronic circuits composed of these operational amplifiers and other electronic components. The time-dependent behavior of the concentration of a protein/molecule can be measured by the voltage at a specific point in the circuits.

The present invention provides a method for molecular dynamics simulation of biomolecules and/or nano-molecular systems comprising:

constructing at least one set of equation representing the molecular dynamics of at least one molecule of the biomolecules and/or the nano-molecular systems;

constructing an electronic circuit representing every set of equation; and

determining molecular dynamics simulation by measuring voltage at two or more connection points of the circuit.

The equations may be linear and/or non-linear second order ordinary differential equations (ODEs). Preferably, non-linear second order ordinary differential equations (ODEs).

The biomolecule and/or nano-molecular system of the invention comprise in general, but not exclusively, amino acids, nucleotides, organic molecules, and/or inorganic molecules.

According to one aspect of the invention, construction plan is disclosed for circuit group of each unit in a biomolecule (amino acid or nucleotide) or a nano-molecular-system to represent all its internal interaction bonded and non-bonded interactions and bonded interactions to its nearest neighbouring amino acid or nucleotide. Each circuit group is composed of at least one of: (1) internal bond stretch, angle bending, torsion, non-bonded units; and (2) bond stretch, angle bending, and torsion units to its two nearest neighbouring amino acid or nucleotide. In a molecular dynamics simulation of a protein or nucleic acid, these circuit groups can be combined together according to the sequence of the protein or nucleic acid. Additional non-bonded units representing inter-residue non-bonded interactions can be added to each circuit group accordingly.

Accordingly, the electronic circuit of the invention may comprise at least one atom-position circuit unit, wherein the atom-position circuit unit represents the position of an atom of a molecule or a molecular system. The atom-position circuit unit comprises at least one atom-atom interaction circuit subunit, the atom-atom interaction circuit subunit representing a sub-unit of atom-atom interactions within a molecule or a molecular system and comprising at least one of: internal bond stretch, angle bending, torsion, non-bonded units; bond stretch, angle bending, and torsion units; between at least two nearest sub-unit of a molecule.

Each atom-atom interaction circuit subunit represents a term in the molecular dynamics equation and comprises at least one of the following: bond stretch x unit, bond stretch y unit, bond stretch z unit, angle bending x type-A unit, angle bending x type-B unit, angle bending y type-A unit, angle bending y

type-B unit, angle bending z type-A unit, angle bending z type-B unit, torsion x type-A unit, torsion x type-B unit, torsion y type-A unit, torsion y type-B unit, torsion z type-A unit, torsion z type-B unit, non-bonded x unit, non-bonded y unit, non-bonded z unit, hydrogen-bond x unit, hydrogen-bond y unit, and hydrogen-bond z unit; and

wherein x, y, and z represent the coordinates of each atom of the molecule, and type-A represents the case of the atom being in the middle-position of an angle bending or torsion connection with other atoms, and type-B represents the case of the atom being in the end-position of an angle bending or torsion connection with other atoms.

According to a particular aspect, all the twenty subunit elements listed above may be used for the construction of the circuit units and electronic circuits of the invention.

In a further aspect of this invention, methods are disclosed for maintaining the voltages in the circuits (xyz coordinates and functions of xyz coordinates) within the allowed range. Scaling factors are applied to x, y, z coordinates and to the molecular dynamics equations so as to keep the x, y, z coordinates and each term in the equations within the allowed range. In addition, resistors and amplifiers can be used to scale-down and scale-up the voltages at certain segment of the circuits. Moreover automatic gain control circuits can also be used to ensure the voltages are within the required range. The automatic gain control circuits are described in "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGram-Hill, 1976.

Accordingly, the method further comprises maintaining the voltage level in the circuit between two fixed voltage values. In particular, x, y and z represent the coordinates of the molecule, and the voltage level of the circuit is maintained between two fixed voltage values by:

applying (non-zero real numbers) scaling factors to the x, y and z coordinates and to the molecular dynamic equation;

applying at least one resistor and/or amplifier at one or more connection point of the circuit, thereby scaling-down or –up the voltage of one or more segment of the circuit; and/or

applying automatic gain control circuits.

According to another aspect of the second embodiment, it is provided a circuit group representing the interaction pattern in the chemical structure of a molecule or a sub-unit of interaction pattern in the chemical structure of a molecule comprising at least one of the following:

a bond stretch connection between each atom pair of the molecule covalently bonded to each other;

an angle bending connection pair between a first atom and other two atoms;

a torsion connection bundle between a first atom and other three atoms; and

a non-bonded connection between each atom pair whose atoms are at least four bonds away from each other.

A circuit unit according to the invention may comprise at least one circuit group as disclosed.

The invention also provides an electronic circuit comprising at least one circuit unit, the circuit unit comprising at least one circuit group, the circuit group representing a sub-unit of interaction pattern in the chemical structure of a molecule and comprising internal bond stretch, angle bending, torsion, non-

bonded units; and/or bond stretch, angle bending, and/or and torsion units; between at least two nearest sub-unit of a molecule.

The electronic circuit represents a term in the molecular dynamic equation, and wherein the circuit unit comprises at least one of the following: bond stretch x unit, bond stretch y unit, bond stretch z unit, angle bending x type-A unit, angle bending x type-B unit, angle bending y type-A unit, angle bending y type-B unit, angle bending z type-A unit, angle bending z type-B unit, torsion x type-A unit, torsion x type-B unit, torsion y type-A unit, torsion y type-B unit, torsion z type-A unit, torsion z type-B unit, non-bonded x unit, non-bonded y unit, non-bonded z unit, hydrogen-bond x unit, hydrogen-bond y unit, and hydrogen-bond z unit; and wherein x, y, and z represent the coordinates of each atom of the molecule, and type-A unit represents the case of the atom being in the middle-position of an angle bending or torsion connection with other atoms, and type-B represents the case of the atom being in the end-position of an angle bending or torsion connection with other atoms.

According to a particular aspect, the invention relates to design plans for the construction of circuit units representing all possible terms in the ODEs for molecular simulation of proteins, nucleic acids and organic molecules. The units are those represented above.

The invention also provides a method for the manufacture of an electronic circuit representing at least one biomolecule and/or nano-molecular system, the electronic circuit comprising at least a unit circuit comprising at least a circuit group, wherein the circuit group represents a sub-unit of interaction pattern in the chemical structure of a molecule or a molecular system comprising:

introducing a bond stretch between each atom of a pair of atoms covalently bonded to each other;

introducing an angle bending connecting pair between a first atom and other two atoms;

introducing a torsion connection bundle between a first atom and other three atoms; and

introducing a non-bonded connection between each atom pair whose atoms are at least four bonds away from each other.

The equations of motion of molecular dynamics simulation of a molecule can be derived from Newton's second law of motion given the potential energy functions, force fields and a starting 3D structure. The potential energy functional for proteins, nucleic acids and organic molecules is given by:

$$V = 1/2\ \Sigma_{bonds} K_r (R - R_{eq})^2 + 1/2\ \Sigma_{angles} K_\theta (\theta - \theta_{eq})^2 + 1/2\ \Sigma_{torsions} V_n [\ 1 + \cos(n\phi - \gamma)\ ] + \Sigma_{non\ bonded} [\ A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6} + q_i q_j/\varepsilon_r r_{ij}] + \Sigma_{H\ bonds} V_H(r) \quad (1)$$

as described in "A second generation force field for simulation of proteins, nucleic acids, and organic molecules", W. D. Cornell et. al., J. Am. Chem. Soc. 117, 5179-5197 (1995); and in "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations", B. R. Brooks, et.al. J. Comp. Chem. 4, 187-217 (1983).

The hydrogen bond potential $V_H(r)$ is different for different force fields. The following is a list of some of the $V_H(r)$ appeared in the literatures:

$$\begin{aligned} V_H(r) &= A/r^{12} - B/r^{6} + q_i q_j/\varepsilon_r r_{ij} && \text{for AMBER (same as non-bonded term)} \\ &= (A/r^{12} - B/r^{10})\cos^m(\theta_{A\text{-}H\text{-}D})\cos^n(\theta_{AA\text{-}A\text{-}H})sw_1(\mathbf{r})sw_2(\theta) && \text{for CHARM} \\ &= V_0 (1 - e^{-a(r-r0)})^2 - V_0 && \text{for Prohofsky/Chen} \end{aligned} \quad (2)$$

All force field parameters for each atom or atom-atom pair in a biomolecule or nano-molecular-machine are given in the literature. For example, AMBER force fields are described in "A second generation force field for simulation of proteins, nucleic acids, and organic molecules", W. D. Cornell et. al., J. Am. Chem. Soc. 117, 5179-5197 (1995). As for organic molecules, all parameters

other than partial charges are available. Partial charges of an organic molecule can be computed from quantum chemistry software such as Gaussian.

From Newton's second law, the atomic position of the i-th atom, in terms of its xyz coordinates (x_i, y_i, z_i) is given by:

$$m_i d^2x_i/dt^2 = -dV/dx_i$$
$$= -\Sigma_{bonds} K_r (R - R_{eq}) B^r_{xi} - \Sigma_{angles} K_\theta (\theta - \theta_{eq}) B^\theta_{xi} + \Sigma_{torsions} nV_n \sin(n\phi-\gamma) B^\phi_{xi}$$
$$- \Sigma_{non\ bonded} [6B_{ij}/r_{ij}^7 - 12A_{ij}/r_{ij}^{13} - q_iq_j/\varepsilon_r r_{ij}](x_i-x_j)/r_{ij} - \Sigma_{H\ bonds} [dV_H(r)/dr](x_i-x_h)/r \quad (3)$$

$$m_i d^2y_i/dt^2 = -dV/dy_i$$
$$= -\Sigma_{bonds} K_r (R - R_{eq}) B^r_{yi} - \Sigma_{angles} K_\theta (\theta - \theta_{eq}) B^\theta_{yi} + \Sigma_{torsions} nV_n \sin(n\phi-\gamma) B^\phi_{yi}$$
$$- \Sigma_{non\ bonded} [6B_{ij}/r_{ij}^7 - 12A_{ij}/r_{ij}^{13} - q_iq_j/\varepsilon_r r_{ij}](y_i-y_j)/r_{ij} - \Sigma_{H\ bonds} [dV_H(r)/dr](y_i-y_h)/r \quad (4)$$

$$m_i d^2z_i/dt^2 = -dV/dz_i$$
$$= -\Sigma_{bonds} K_r (R - R_{eq}) B^r_{zi} - \Sigma_{angles} K_\theta (\theta - \theta_{eq}) B^\theta_{zi} +/ \Sigma_{torsions} nV_n \sin(n\phi-\gamma) B^\phi_{zi}$$
$$- \Sigma_{non\ bonded} [6B_{ij}/r_{ij}^7 - 12A_{ij}/r_{ij}^{13} - q_iq_j/\varepsilon_r r_{ij}](z_i-z_j)/r_{ij} - \Sigma_{H\ bonds} [dV_H(r)/dr](z_i-z_h)/r \quad (5)$$

where m_i is the mass of the i-th atom; B^r_{xi}, B^r_{yi}, B^r_{zi} are the x, y z components of the B-matrix for bond stretching; B^θ_{xi}, B^θ_{yi}, B^θ_{zi} are the x, y, z components for the B-matrix for bond angle bending; B^ϕ_{xi}, B^ϕ_{yi}, B^ϕ_{zi} are the x, y, z components for the B-matrix for torsion. As described in "Vibrational states", S. Califano. John Wiley & Sons, New York (1976), these B-matrix elements are given by:

$$B^r_{xi} = -A_{ij} \quad (6)$$

$$B^r_{yi} = -B_{ij} \quad (7)$$

$$B^r_{zi} = -C_{ij} \quad (8)$$

$$B^\theta_{xi} = Ea = -(Eb+Ec) \quad \text{if I-th atom is in the middle}$$
$$= Eb \quad \text{if I-th atom is at end} \quad (9)$$

$$B^\theta_{yi} = Fa = -(Fb+Fc) \quad \text{if I-th atom is in the middle}$$
$$= Fb \quad \text{if I-th atom is at end} \quad (10)$$

$$B^\theta_{zi} = Ga = -(Gb+Gc) \quad \text{if I-th atom is in the middle}$$
$$= Gb \quad \text{if I-th atom is at end} \quad (11)$$

$$B^{\phi}_{xi} = Ua = 1/r_{ij}p_a^2 + q_c s_5/r_{ik}p_c^2 - q_a s_6/r_{ik}p_a^2 \quad \text{if I-th atom is in the middle}$$
$$= Ub = -\, s_6/r_{ij}p_a^2 \quad \text{if I-th atom is at end} \qquad (12)$$

$$B^{\phi}_{yi} = Va = 1/r_{ij}p_a^2 + q_c s_3/r_{ik}p_c^2 - q_a s_4/r_{ik}p_a^2 \quad \text{if I-th atom is in the middle}$$
$$= Vb = -\, s_4/r_{ij}p_a^2 \quad \text{if I-th atom is at end} \qquad (13)$$

$$B^{\phi}_{xi} = Wa = 1/r_{ij}p_a^2 + q_c s_1/r_{ik}p_c^2 - q_a s_2/r_{ik}p_a^2 \quad \text{if I-th atom is in the middle}$$
$$= Wb = -\, s_2/r_{ij}p_a^2 \quad \text{if I-th atom is at end} \qquad (14)$$

where r_{ij}, A_{ij}, B_{ij}, C_{ij}, q_a, q_c, p_a, p_c, Eb, Ec, Fb, Fc, Gb, Gc, s_1, s_2, s_3, s_4, s_5, s_6 are given by:

$$r_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{1/2} \qquad (15)$$

$$A_{ij} = (x_j - x_i)/r_{ij} \qquad (16)$$

$$B_{ij} = (y_j - y_i)/r_{ij} \qquad (17)$$

$$C_{ij} = (z_j - z_i)/r_{ij} \qquad (18)$$

$$q_a = \cos\theta_a = A_{ij}A_{ik} + B_{ij}B_{ik} + C_{ij}C_{ik} \qquad (19)$$

$$q_c = \cos\theta_c = -(A_{ik}A_{kl} + B_{ik}B_{kl} + C_{ik}C_{kl}) \qquad (20)$$

$$p_a = \sin\theta_a = (1 - q_a)^{1/2} \qquad (21)$$

$$p_c = \sin\theta_c = (1 - q_c)^{1/2} \qquad (22)$$

$$Eb = (A_{ij}q_a - A_{ik})/r_{ij}p_a \qquad (23)$$

$$Ec = (A_{ik}q_a - A_{ij})/r_{ik}p_a \qquad (24)$$

$$Fb = (B_{ij}q_a - B_{ik})/r_{ij}p_a \qquad (25)$$

$$Fc = (B_{ik}q_a - B_{ij})/r_{ik}p_a \qquad (26)$$

$$Gb = (C_{ij}q_a - C_{ik})/r_{ij}p_a \qquad (27)$$

$$Gc = (C_{ik}q_a - C_{ij})/r_{ik}p_a \qquad (28)$$

$$S_1 = A_{ik}B_{kl} - B_{ik}A_{kl} \qquad (29)$$

$$S_2 = A_{ik}B_{ij} - B_{ik}A_{ij} \qquad (30)$$

$$S_3 = A_{kl}C_{ik} - C_{kl}A_{ik} \qquad (31)$$

$$S_4 = A_{ij}C_{ik} - C_{ij}A_{ik} \qquad (32)$$

$$S_5 = B_{ik}C_{kl} - C_{ik}B_{kl} \quad (33)$$

$$S_6 = B_{ik}C_{ij} - C_{ij}B_{ik} \quad (34)$$

This set of nonlinear ODEs contain terms of r_{ij}, $(x_i-x_j)\times(x_k-x_j)$, $(x_i-x_j)/r_{ij}$, $\sin(n\phi-\gamma)$ etc., which can be represented by a corresponding set of electronic analog circuits composed of known operational amplifiers, function circuits and other circuit components.

Referring now to Figure 10, Figure 11, and Figure 12, there are disclosed design plans for construction of electronic circuits representing molecular dynamics simulation equations (3), (4) and (5) for each atom in a protein, nucleic acid, or organic molecule respectively. For each equation, its circuit starts with a point x_i or y_i or z_i, which is connected to a switch directed to a voltage source representing the initial coordinate. This point is also connected to other parts of the circuit or other circuits in a combination of one or more of the seven cases as illustrated in Figure 10, Figure 11, and Figure 12. In a simulation, the time-dependent x or y or z coordinate of the i-th atom can be by measurement of voltages at x_i or y_i or z_i .

Construction Plan for the Specially Designed Circuit Groups Representing All the Internal Terms and Bonded terms with nearest neighbors in the ODEs of a sub-unit of a biomolecule or in a nano-molecular-machine

Referring now to Figure 13 to Figure 15, there are disclosed design plans for construction of circuit group for each illustrative sub-unit, amino acid, of a biomolecule, protein, to represent all its internal terms and bonded terms to its nearest neighboring subunits, amino acids.

Referring now to Figure 17 to Figure 19, there are disclosed design plans for construction of circuit group for each sub-unit of an illustrative sub-unit of a

nano-molecular-machine in Figure 16 to represent all its internal terms and bonded terms to its nearest neighboring subunits.

In these figures, each box represents the x, y and z circuit sections of each atom of an amino acid or nucleotide. An arrow represents a connection between two atoms such that the output voltage x_j, y_j, or z_j from the x, y or z circuit section of an atom is directed to a bond-stretch, angle bending or torsion unit in the corresponding section of the atom to which the arrow is pointed. For clarity purpose, the connection profile for bond stretch, angle bending, and torsion is displayed in separate Figures. In reality, these should be put together in the same circuit group configuration. Figure 13 and Figure 17 illustrates connections to bond stretch units of each circuit section, Figure 14 and Figure 18 illustrates connections to angle bending units of each circuit section, and Figure 15 and Figure 19 illustrates connections to torsion units of each circuit section.

A circuit group is in general constructed by the following rules:

From the chemical structure of an amino acid or nucleotide:

A bond stretch connection is introduced to each atom pair covalently bonded to each other;

An angle bending connection pair are established from atom A to two other atoms B and C, in cases that they are linearly bonded as A-B-C or B-A-C or B-C-A etc.;

A torsion connection bundle are constructed from atom A to three other atoms B, C and D, in cases that they are linearly bonded as A-B-C-D or B-A-C-D or B-C-A-D etc.;

A non-bonded connection is introduced to each atom pair whose atoms are at least 4 bonds away from each other.

Construction Plan for the Specially Designed Circuit Units Representing Different Terms in an ODE For Molecular Dynamics Simulation

Referring now to Figure 20 to Figure 40, there are disclosed design plans for construction of special circuit units representing all possible terms of an ordinary differential equation associated with a biological or chemical reaction pathway. The operational amplifiers and random noise generator in these units are all described in the literature. The inverting amplifier and noninverting amplifier are described in "Electronic circuit analysis and design, 2nd edition" D.A. Neamen, McGraw-Hill, 2001. The multiplier is described in "Fundamentals of linear circuits" T.L. Floyd, Macmillan Publishing Co. 1992. The exponentiator is described in "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGram-Hill, 1976. The sine function amplifier is described in "Modern operational circuit design" J. I. Smith, Wiley-Interscience, 1971. The $\cos^{-1}$ function amplifier can be generated in different ways. One example is the use of circuits of polynomials and another example is the use of circuits of non-integer exponent as described in "Function circuits: Design and applications" Y.J. Wong and W.E. Ott, McGram-Hill, 1976.

Some circuit units in Figure 10 to Figure 40 use one or more of the following specially introduced circuit sun-units: rABC unit, AqA unit, pq unit, Ua unit, ss unit, si unit, ϕ unit, $\sin\phi$ unit. These sub-units are given in Figure 41 to Figure 44.

Methods for Maintaining the Voltages in the Circuits Within Allowed Range

In large proteins, the distance between two atoms can be larger than a few hundred angstroms. Thus the range of the xyz coordinates in these proteins

may exceed the allowed voltages in a circuit. In addition, some terms in the molecular dynamics simulation equations can become very large if two atoms get too close or too far apart from each other, thereby causing the voltages at certain parts of the circuits to exceed the allowed range.

These problems can be solved by maintaining the voltage level in the circuit between two fixed voltage values, by using at least one of the following methods:

(1) Applying scaling factors to x, y, z coordinates and to the molecular dynamics equations so as to keep the x, y, z coordinates and each term in the equations within the allowed range.

This can be accomplished by multiplying one or two appropriately selected small numbers (non-zero real numbers) δ or $\delta\varepsilon$ to the Equation 3, Equation 4, and Equation 5.

(2) Resistors are used to scale-down the voltages at the input end of a segment of the circuits and amplifiers are used at the output end of the segment to scale-up the voltages to the proper level.

(3) Automatic gain control circuits can be added to the circuits to ensure the voltages are within the required range.

Illustrative example: Electronic protein emulator

Structural optimization and molecular dynamics of a protein can be described by a set of ODEs. An illustrative amino acid is governed by the collection of ODEs of each constituent atom (Equation 3-5). The circuit for each atom is

illustrated in Figure 10 to Figure 12. The circuit group for all the atoms of an amino acid is illustrated in Figure 13 to Figure 15.

Illustrative example: Electronic nano-molecular-system emulator

There have been efforts for developing nano-molecular-systems. One example is a mimic of an enzyme ribonuclease shown in Figure 16.

Structural optimization and molecular dynamics of such a nano-molecular-machine can be described by a set of ODEs. An illustrative sub-unit of cyclodextrin shown in Figure 16(a) is governed by the collection of ODEs of each constituent atom (Equation 3-5). The circuit for each atom is illustrated in Figure 10 to Figure 12. The circuit group for all the atoms of a sub-unit is illustrated in Figure 17 to Figure 19.